NAT'L INST. OF STAND & TECH R.I.C.
A11103 709213

NIST
PUBLICATIONS

NISTIR 4715

KINEMATIC CALIBRATION OF AN ACTIVE CAMERA SYSTEM

G. S. Young

Robotics Laboratory
Department of Mechanical Engineering
University of Maryland
College Park, MD 20742

T. H. Hong

Department of Computer Science and
Information Services
The American University
Washington, DC 20016

M. Herman

Sensory Intelligence Group
Robot Systems Division

J. C. S. Yang

Robotics Laboratory
Department of Mechanical Engineering
University of Maryland
College Park, MD 20742

U.S. DEPARTMENT OF COMMERCE
National Institute of Standards
and Technology
Robot Systems Division
Sensory Intelligence Group
Bldg. 220, Room B124
Gaithersburg, MD 20899

U.S. DEPARTMENT OF COMMERCE
Robert A. Mosbacher, Secretary
NATIONAL INSTITUTE OF STANDARDS
AND TECHNOLOGY
John W. Lyons, Director

QC
100
.U56
#4715
1991
C.2

NIST

NATIONAL INSTITUTE OF STANDARDS &
TECHNOLOGY
Research Information Center
Gaithersburg, MD 20899

NISTIR 4715

KINEMATIC CALIBRATION OF AN ACTIVE CAMERA SYSTEM

G. S. Young

Robotics Laboratory
Department of Mechanical Engineering
University of Maryland
College Park, MD 20742

T. H. Hong

Department of Computer Science and
Information Services
The American University
Washington, DC 20016

M. Herman

Sensory Intelligence Group
Robot Systems Division

J. C. S. Yang

Robotics Laboratory
Department of Mechanical Engineering
University of Maryland
College Park, MD 20742

U.S. DEPARTMENT OF COMMERCE
National Institute of Standards
and Technology
Robot Systems Division
Sensory Intelligence Group
Bldg. 220, Room B124
Gaithersburg, MD 20899

November 1991

U.S. DEPARTMENT OF COMMERCE
Robert A. Mosbacher, Secretary

NATIONAL INSTITUTE OF STANDARDS
AND TECHNOLOGY
John W. Lyons, Director

ABSTRACT

Perceptual activity for exploration, probing and searching is very important in computer vision. For the purpose of intelligently controlling the sensor's motion and parameters for different sensing strategies and designated tasks in perceptual activity, an active camera system is often used. An active camera system is also usually used for vision-based guidance. In order to position the camera system accurately and to obtain the relation between the camera and the manipulator, the active camera system must be calibrated.

In this paper, we introduce a modified Denavit-Hartenburg kinematic model and develop a new technique to calibrate an active camera system. The manipulator, camera-to-manipulator, camera, and base-to-world calibrations are all included. The method employs four ideas : (1) The camera poses, joint poses, and link frames calibrated are all related to the world frame, therefore the camera-to-manipulator and base-to-world calibration is very straightforward; (2) The joint poses are calibrated separately; (3) The manipulator motion is obtained from the camera poses; (4) Once the joint poses are obtained, the link frames can be defined for any kinematic model. In consequence, the whole procedure is simple, flexible, accurate, and efficient. Two experiments are performed to verify the accuracy of the new technique.

1 INTRODUCTION

1.1 Motivation

Many researchers in computer vision have recognized the importance of perceptual activity for exploration, probing, and searching [5] [2]. Passive sensors such as CCD (Charge Coupled Device) cameras mounted on an active camera system are often used in perceptual activity. By intelligently controlling the sensor's motion and parameters for different sensing strategies and designated tasks, perceptual activity can be accomplished. Many successful applications of passive cameras used in perceptual activity have been reported. In [3], by intelligently controlling the camera motion, solutions to the shape from shading, contour, and texture problems become well-conditioned and unique. Reference [4] discusses how using a purposive and qualitative approach robustly solves many navigational problems without reconstructing the scene. Camera fixation (defined as actively controlling the camera motion so that a fixed point in space is always imaged to the same position in the image plane) makes image interpretation easier[17]. Using two types of constrained camera motion, the depth of objects in bins can be inferred[24]. Active camera systems are used not only for perceptual activity in computer vision but also in robotics. For example, an active camera system is often used for vision-base guidance[25].

In this paper, we consider an active camera system which has one or more cameras rigidly mounted on a manipulator, such as a robot arm or a positioning table. The manipulator may be modeled as a set of n moving rigid bodies connected in a serial chain by any combination of revolute and/or prismatic joints. These moving bodies are called links. Each link has one degree of freedom (DOF).

A vital issue for active camera systems is the ability to accurately know the state of the system and how this state was achieved. Very often, the positioning inaccuracy of the manipulator may occur due to manufacturing tolerances and the assembly process. For example, in our vision laboratory a camera is mounted on a two DOF positioning table, with one prismatic joint and one revolute joint. Ideally, for convenience and simple computation, we want the two joint axes of the positioning table to intersect and to be perpendicular to each other. Also, we want the camera center to be located at the center of the rotation axis and the initial viewing direction of the camera to be parallel to one of the joint axes. In practice, it is very difficult to accurately set up the above ideal configuration. The camera pose (position and orientation) is unknown. Furthermore, the manipulator motion is typically represented in terms of the manipulator base frame, while the visual data is often expressed with respect to the camera frame. To relate the camera and the manipulator, the transformation between these two frames must be determined. Unfortunately, it is very hard to directly and accurately measure the transformation. Therefore, calibration of the active camera system is required to improve the positioning accuracy and to determine the transformation between the camera and the manipulator.

Due to the importance of calibration, the goal of this research is to develop an accurate and efficient method to calibrate the active camera system. A unified calibration of an active

camera system (see figure 1) includes four major parts:
(1) The manipulator calibration ($H_m^B(\mathrm{V})$, a 4x4 transformation matrix and function of the commanded joint readings V) that locates the end frame relative to the base frame.
(2) The camera-to-manipulator calibration (H_c^m, a 4x4 matrix) that finds the transformation between the camera and the manipulator.
(3) The camera calibration that characterizes the inherent properties of the camera and optics and determines the camera pose ($H_W^c(\mathrm{V})$, a 4x4 matrix and function of the commanded joint readings V) between the world frame and the camera frame.
(4) The base-to-world calibration (H_B^W, a 4x4 matrix) that defines the transformation between the base frame and the world frame.
In figure 1, $H_c^B(\mathrm{V})$ describes the transformation between the camera frame and the base frame and is a function of the commanded joint readings V. All the above transformations are expressed in solid lines or dash lines in figure 1. The solid lines represent the fixed relation between two frames, while the dash lines means that the relation is a function of V.

1.2 Previous Work

In previous work for the manipulator calibration (H_m^B), the algorithms differ from one another in three ways: the kinematic model representation, the manipulator parameter identification, and the measurement devices used. A survey may be found in [13]. The selection of the kinematic model varies in different approaches. Based on the definition of the coordinate transformation, four, five, six, or even more parameters may be used. Most of the methods for identifying the manipulator parameters fall into two categories. One is to formulate the problem as a set of nonlinear optimization equations and solve for all joints parameters at one time [16]. The other is to use analytical techniques to solve the parameters for each individual link [8] [27] [15]. As to the choice of the instrumentation for calibration, most of the methods require expensive devices such as laser scanning systems, laser tracking interferometer systems, selspots, or theodolites. On the other hand, Tsai's method [15] employs a general purpose, off-the-shelf common CCD camera and a calibration target block.

One method of achieving a camera-to-manipulator (sometimes called hand-eye, head-eye, etc.) calibration (H_c^m) is to integrate it with the aforementioned manipulator calibration, then formulate a large scale nonlinear optimization problem and solve the parameters for both parts simultaneously [16] [6]. This approach is often time-consuming, and it needs a good initial guess for convergence. Another approach used by [18] [10] [23] [9] is to treat the camera-to-manipulator calibration (H_c^m) alone, and solve it from the equation

$$A\, H_c^m \;=\; H_c^m\, B$$

Here A, B, and H_c^m are all 4x4 matrices. A is the change in the manipulator position and B is the resulting camera displacement. Note that in order to apply this method A and B must be given to calculate H_c^m. This solution is non-trivial.

For the topic of camera calibration, a detailed survey may be found in [22]. A fast and accurate method of obtaining the intrinsic and extrinsic parameters which relate the 2D camera image and the 3D scene is described in [20]. The implementation of Tsai's camera calibration algorithm can also be found in [19].

Calibration, treated as a unified problem, has been discussed in several papers. Puskorius [16] integrated manipulator calibration and camera-to-manipulator calibration into one large scale nonlinear optimization problem. In this approach, an accurate initial guess is often required and a large number of parameters must be solved. The trio method developed by Tsai and Lenz, which includes manipulator calibration, hand-eye calibration, and camera calibration has been a major contribution to unified calibration [21]. This approach is generally good, but there are limitations. Instead of using a general kinematic model for the manipulator, this method focuses only on Cartesian robots. In order to calibrate the rotation axis, some linear joints may also be exercised to keep the calibration block within the field of view. Although there are two elements (the orientation and the position of the rotation axis) to be found, only the orientation can be calculated by this method. The position is still unknown. The hand-eye calibration is currently tedious, but there is room for improvement. The work by Bennett et al [6] is also in the category of unified calibration. In this approach, no external measurement device is needed. However, similar to Puskorius'es approach, an accurate initial guess is often required and a large number of parameters must be solved.

1.3 Our Approach

Our approach has the following features:

(1) It is a unified calibration method: The manipulator, camera-to-manipulator, camera, and base-to-world calibrations are all included.

(2) The camera poses, joint poses, and link frames are all related to the world frame. Therefore the camera-to-manipulator and base-to-world calibration is very straightforward.

(3) Most importantly, the camera-to-manipulator calibration as well as the base-to-world calibration are the by-products of the manipulator calibration. The procedure of solving A $H_c^m = H_c^m$ B is no longer necessary.

(4) We modify the Denavit-Hartenburg kinematic model and introduce two parameters, scale factor and zero-offset, to correct the joint variable values used from the joint encoder readings.

(5) Since the method calibrates the pose relative to the world frame for each individual joint, only the form of one-equation-solving-one-parameter is used. Therefore, the computation is simple. It is also flexible because any proper starting position can be used for each joint calibration, and the sequence in which joints are calibrated does not matter.

(6) The method is good for a manipulator with any type of kinematic model. Since each joint axis pose solved by our approach is expressed in reference to the world frame, once the joint pose is obtained, the individual link parameters can be extracted and the link frame can be defined for any kinematic model.

(7) No expensive measurement device is needed. Only a CCD camera and a calibration target are used.

(8) The kinematic camera pose can be written as a function of the commanded joint readings. The advantage of this is that whenever camera motion is induced by the manipulator, we can substitute the new joint readings into the existing camera pose function, and the new camera pose can be obtained right away. There is no need to calibrate the camera pose again.

1.4 Outline of This Paper

In the following sections, some backgound for our derivations is first introduced, then the procedures and equations for our unified calibration method are elaborated. Following that, the experiments for verifying this method are discussed and the results are presented. Finally, some conclusions are drawn for our method.

2 THEORY

2.1 Background

The background below is useful to understand our mathematical derivations.

- **Coordinate frames**

All coordinate frames defined here are Cartesian coordinate frames, as follows:

c: The camera coordinate frame (figure 2) is attached to the camera, with the origin at the camera center, the Z axis coinciding with the optical axis, and the X, Y axes parallel to the image x and y axes, respectively. Note that the camera frame is a moving coordinate frame, depending on where the camera is.

W: The world coordinate frame (figure 3) is fixed and arbitrarily selected. The coordinate of each target point on the calibration plate is known a priori with respect to this frame.

L: The link coordinate frame i of the manipulator is attached to link i (figure 4). Let Z_i and Z_{i+1} lie along the joint axes i and i+1 at links i and i+1, respectively. X_i is chosen as the unit vector perpendicular to both Z_i and Z_{i+1}, and Y_i is found by taking the cross product of Z_i and X_i. Note that the link frame i is a moving coordinate frame. The pose of X_i depends on the commanded joint readings at joints 1 through i while the pose of Z_i depends on the readings at joint 1 through i-1.

B: The manipulator base coordinate frame is regarded as the reference frame of the manipulator fixed in space.

- **Transformations**

Let us introduce three interpretations of the homogeneous transform H, a 4x4 matrix containing orientation and position information[11]:

$$H = \begin{bmatrix} R & T \\ 0\,0\,0 & 1 \end{bmatrix} \tag{1}$$

where R is a 3x3 rotation matrix, and T is a 3x1 translation vector

(1) Description of a frame

H_B^A describes the frame B relative to the frame A.

$$H_B^A = \begin{bmatrix} X_B^A & Y_B^A & Z_B^A & O_B^A \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} R_B^A & T_B^A \\ 0\,0\,0 & 1 \end{bmatrix} \tag{2}$$

where X_B^A,Y_B^A, and Z_B^A are 3x1 unit vectors representing the coordinate axes X,Y, and Z of frame B, respectively, and O_B^A is the position vector of the origin of frame B relative to frame A, R_B^A is a 3x3 rotation matrix of frame B relative to A, and T_B^A is a 3x1 translation vector of frame B relative to A.

(2) Transform mapping: changing descriptions from frame B to frame A

In figure 5, we have a physical point P in space. P^A and P^B are the 3x1 position vectors of point P in frames A and B, respectively.

We can calculate P^A from

$$\begin{Bmatrix} P^A \\ 1 \end{Bmatrix} = H_B^A * \begin{Bmatrix} P^B \\ 1 \end{Bmatrix} \tag{3}$$

(3) Transform operator: only one reference frame A involved

The operator $H_{1,2}^A$ rotates and translates a coordinate frame H_1^A attached to a rigid body into a new pose H_2^A (see figure 6):

$$H_2^A = H_{1,2}^A * H_1^A \tag{4}$$

The following are some basic formulas for transformations:

$*$ Compound transformations

$$H_C^A = H_B^A * H_C^B \tag{5}$$

or

$$\begin{bmatrix} R_C^A & T_C^A \\ 0\,0\,0 & 1 \end{bmatrix} = \begin{bmatrix} R_B^A * R_C^B & R_B^A * T_C^B + T_B^A \\ 0\,0\,0 & 1 \end{bmatrix} \tag{6}$$

Note that $R_B^A * R_C^B \neq R_C^B * R_B^A$

$*$ Inverting a transform

$$H_A^B = \left(H_B^A\right)^{-1} \tag{7}$$

or

$$\begin{bmatrix} R_A^B & T_A^B \\ 0\,0\,0 & 1 \end{bmatrix} = \begin{bmatrix} \left(R_B^A\right)^t & -\left(R_B^A\right)^t * T_B^A \\ 0\,0\,0 & 1 \end{bmatrix} \tag{8}$$

where -1 and t denote inverse and transpose of a matrix, respectively.

2.2 Models

Two models that will be used in our derivation are described next:

Manipulator model

Consider a general serial link manipulator of arbitrary order and with a combination of revolute and prismatic joints. In order to deal with the complex geometry of a manipulator, we affix a coordinate frame to each link and then describe the relationship between these coordinate frames. A general method for assigning link coordinate frames and parameters to a manipulator was developed in [12] (see figure 7). The location and orientation of the i-th link coordinate frame relative to the i-1 th link coordinate frame can be represented as a 4x4 homogeneous transformation matrix. The Denavit-Hartenburg link transformation is obtained from [11]

$$H_i^{i-1} = \begin{bmatrix} c\theta_i & -s\theta_i & 0 & a_{i-1} \\ s\theta_i c\alpha_{i-1} & c\theta_i c\alpha_{i-1} & -s\alpha_{i-1} & -s\alpha_{i-1} d_i \\ s\theta_i s\alpha_{i-1} & c\theta_i s\alpha_{i-1} & c\alpha_{i-1} & c\alpha_{i-1} d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{9}$$

where s and c refer to the sine and cosine function, a is the link length, α is the link twist angle, d is the link offset, and θ is the joint angle.

In order to correct for the joint variable values used from the commanded joint readings, we modify the D-H model and introduce two parameters, scale factor and zero-offset, represented as follows:

$$\theta_i = \theta_i^* + s_i \cdot v_i \;, \qquad d_i = d_i^* \;\; for \;\; a \;\; revolute \;\; joint \tag{10}$$

and

$$d_i = d_i^* + s_i \cdot v_i \;, \qquad \theta_i = \theta_i^* \;\; for \;\; a \;\; prismatic \;\; joint \tag{11}$$

where s_i is the scale factor, v_i is the commanded joint reading, $\theta_i{}^*$ is the zero-offset of the joint angle, and $d_i{}^*$ is the zero-offset of the link. The scale factor s_i is often required because the commanded joint reading v_i must be converted into the real value θ_i for the revolute joint or d_i for the prismatic joint. On the other hand, the use of a zero offset $\theta_i{}^*$ or d_i^* can adjust θ_i or d_i to correct value. Thus, due to the modification of equations (10) and (11), H_i^{i-1} in equation (9) contains the constant parameters a_{i-1}, α_{i-1}, $d_i{}^*$, $\theta_i{}^*$, s_i and one variable v_i.

For convenience, we choose the manipulator base frame (called B or 0) to coincide with frame 1 when the commanded joint reading v_1 of link 1 is zero. Thus, we will always have

$$a_0 = \alpha_0 = d_1^* = \theta_1^* = 0 \tag{12}$$

Similarily, for the frame of the last link m, the direction of X_m is chosen to be aligned with X_{m-1} when the commanded joint reading v_m of link m is zero. Thus, we will always have

$$d_m^* = \theta_m^* = a_m = \alpha_m = 0 \tag{13}$$

In consequence, the location of the last link m is related to the base frame B or 0 by a sequence of D-H transformations as follows:

$$H_m^0 = H_1^0 * H_2^1 * \cdots * H_m^{m-1} \tag{14}$$

Camera model

A refined camera model from Tsai [20] is described in figure 8 and used to transform the 3-D object coordinates(X_W, Y_W, Z_W) to the 2-D computer image coordinates (x_f, y_f) in pixels using the following steps.

Step 1: Transform the mapping from the world coordinates (X_W, Y_W, Z_W) to the camera coordinates (Xc,Yc,Zc)

$$\begin{Bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{Bmatrix} = H_W^c * \begin{Bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{Bmatrix} = \begin{bmatrix} R_W^c & T_W^c \\ 0\,0\,0 & 1 \end{bmatrix} * \begin{Bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{Bmatrix} \tag{15}$$

where H_W^c, R_W^c, and T_W^c are a 4x4 transform matrix, a 3x3 rotation matrix, and a 3x1 translation vector, respectively, from the world frame W to the camera frame c and

$$R_W^c \equiv \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix}, \quad T_W^c \equiv \begin{Bmatrix} T_x \\ T_y \\ T_z \end{Bmatrix} \tag{16}$$

Step 2: Perspective projection with pin hole geometry

$$x_u = f \cdot \frac{X_c}{Z_c} \tag{17}$$

$$y_u = f \cdot \frac{Y_c}{Z_c} \tag{18}$$

where $(x_u,\ y_u)$ is the undistorted 2-D image coordinate and f is the effective focal length.

Step 3: Radial lens distortion

$$x_d = 2x_u/D, \qquad y_d = 2y_u/D \tag{19}$$

with

$$D = 1 + (1 - 4k_1(x_u^2 + y_u^2))^{1/2} \tag{20}$$

where (x_d, y_d) is the distorted 2-D image coordinate and k_1 is the distortion coefficient.

Step 4: Relationship between the computer image coordinate $(x_f,\ y_f)$ and the distorted image coordinate (x_d, y_d)

$$x_f = s_x \cdot x_d + C_x \tag{21}$$

$$y_f = s_y \cdot y_d + C_y \tag{22}$$

where $(C_x,\ C_y)$ is the computer image coordinate for the origin in the image plane and s_x, s_y are the scale factors in x and y respectively.

The parameters of the camera model may be decoupled into the intrinsic and extrinsic parameters. The intrinsic parameters characterize the inherent properties and are the effective focal length f (lens principal distance), lens distortion coefficient k_1, image scale factors s_x and s_y (relationship of the computer image coordinates in pixels to the true image coordinates), and image center (C_x, C_y) (the point where the optical axis intersects with the image plane). The extrinsic parameters are the position T_W^c and orientation R_W^c of the world reference frame relative to the camera frame.

2.3 Derivation

A unified calibration includes the manipulator, camera-to-manipulator, camera, and base-to-world calibrations. In our approach, most of the effort is concentrated on the manipulator calibration. After the manipulator is calibrated, the camera-to-manipulator and base-to-world calibration can be considered the by-products of the manipulator calibration. To achieve camera calibration, we adopt Tsai's algorithm [20].

In this section, we first describe two procedures: the procedure for solving the transform matrix H_m^B and the procedure for extending H_m^B to H_c^B and H_W^c . Note that H_m^B, H_c^B, and H_W^c will be represented as functions of commanded joint readings V. In the last part of this section, we briefly describe the approach used for Tsai's camera calibration.

Procedure for solving H_m^B(V)

The modified D-H model previously described in equations (9), (10), and (11) is used in the following procedure.

Step 1: For each joint axis $i = 1, \cdots, m$, calibrate its poses at configuration j and the scale factor s_i.
The calibrated joint pose relative to the world frame, $Z_i^W(V_j)$ and $D_i^W(V_j)$, is a function of V where $Z_i^W(V_j)$ is the orientation of the joint axis i, $D_i^W(V_j)$ is a reference point lying on the joint axis, and V_j is a vector $(v_1, v_2, \cdots, v_m)$ of joint readings, storing all the commanded values of the joint positions at configuration j. Note that the calibrations of the joints are performed independently of one another, and the sequence in which they are calibrated is arbitrary.

For each joint i, do the following:

1-1 Move the camera to a *proper* starting location. (By *proper*, we mean that the camera poses at the starting location by moving one or more joints and at the next locations which are due to the movements about the joint axis i can be estimated without difficulty.) Record all commanded joint readings V at the starting location (configuration j=i). Then, estimate the camera pose relative to the world frame at the starting configuration, called H_W^{c1}. Many algorithms are available for estimation of camera pose [14] [1] [20]. Here we use Tsai's *camera calibration algorithm* [20] to estimate camera pose. Note that the starting position for the calibration of each joint does not have to be the same as for other joints.

1-2 Exercise only joint i. Record the difference of two commanded joint readings between the starting and current manipulator configurations, called Δv_i, then estimate the current camera pose, called H_W^{c2}. For higher accuracy, we may repeat **1-2** and obtain many Δv_i and H_W^{c2}.

1-3 Calculate the rigid body motion of link i.
Because the camera is rigidly mounted on the manipulator and only joint i is exercised in **1-2**, the rigid body motion of link i is identical to the motion of the camera. To calculate the movement (rigid body motion) of the camera between c1 and c2, we know

$$H_{c_2}^W = H_{1,2}^W * H_{c_1}^W \tag{23}$$

and therefore

$$H_{1,2}^W = (H_W^{c_2})^{-1} * H_W^{c_1} \equiv \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{24}$$

Equation (24) represents the movement of link i.

1-4 Calculate Z_i^W, D_i^W, and s_i.
For a revolute joint:
Since the motion from H_W^{c1} to H_W^{c2} only involves a rotation about the joint axis i, the screw displacement equations [7] can be applied as follows:
The rotation angle ϕ and the orientation of the joint axis Z_i^W relative to the world coordinate frame can be expressed as

$$\phi_i = \cos^{-1}\left(\frac{m_{11} + m_{22} + m_{33} - 1}{2}\right) \tag{25}$$

and

$$Z_i^W = \begin{Bmatrix} \frac{m_{32} - m_{23}}{2\sin\phi_i} \\ \\ \frac{m_{13} - m_{31}}{2\sin\phi_i} \\ \\ \frac{m_{21} - m_{12}}{2\sin\phi_i} \end{Bmatrix} \tag{26}$$

where the m_{ij} are from equation (24).
One point D_i^W on the joint axis with shortest distance from the origin of the world frame is selected to be a reference point and is described below.

$$D_i^W = \frac{1}{2}\left\{O_{c1}^W + O_{c2}^W + \left[Z_i^W \times \left(O_{c1}^W - O_{c2}^W\right) / \tan\left(\frac{\phi_i}{2}\right)\right] - \left[Z_i^W \cdot \left(O_{c1}^W - O_{c2}^W\right) Z_i^W\right]\right\} \tag{27}$$

where $O_{ci}^W = {O_W^{ci}}^{-1}$ represents the position vector of the origin of camera frame c_i relative to the world frame W.
The scale factor s_i can be obtained as follows:

$$s_i = \frac{\phi_i}{\Delta v_i} \tag{28}$$

For a prismatic joint:

The orientation Z_i^W of the joint axis in terms of the world frame can be obtained from

$$Z_i^W = \frac{O_{c2}^W - O_{c1}^W}{\left|O_{c2}^W - O_{c1}^W\right|} \tag{29}$$

One point D_i^W is arbitrarily chosen as a reference point. The scale factor s_i is

$$s_i = \frac{\left|O_{c2}^W - O_{c1}^W\right|}{\Delta v_i} \tag{30}$$

We may take the averages for the calibrated pose of the joint axis i and scale factor s_i by repeating **1-2** to obtain more H_W^{c2} and Δv_i values. The average Z_i^W and s_i will be :

$$Av.\ Z_i^W = \frac{\sum_1^l Z_i^W}{l}$$

$$Av.\ s_i = \frac{\sum_1^l s_i}{l} \tag{31}$$

where l is the number of data sets.

Step 2: Extraction of the constant link parameters a_i, α_i, d_i^*, θ_i^*
In order to find the constant link parameters, Z_i^W and Z_{i+1}^W calibrated in the same configuration are needed. Since the procedure of manipulator calibration is flexible, there may be different configurations for the two neighboring joint axes due to the different starting calibration locations. Therefore, it may be necessary to transform $Z_i^W(V_{j=i})$ into $Z_i^W(V_{j=i+1})$. After extracting $Z_i^W(V_{j=i+1})$ and $Z_{i+1}^W(V_{j=i+1})$, the next steps involve determining the line which is perpendicular to both axes (figure 9), computing the intersection of this line with each axis, obtaining the constant link parameters, and describing the link coordinate frame. The above procedure is elaborated below for the cases $i = 1$ and $2 \leq i < m$.

Case 1: $i = 1$

$*$ Transform both axes 1 and 2 to the same starting configuration j=2:
The pose of joint axes 2 at configuration j=2 has already been found in step 1. Joint axis 1 is always fixed relative to the base, therefore we have

$$Z_1^W(V_{j=2}) = Z_1^W(V_{j=1}) \tag{32}$$

and

$$D_1^W(V_{j=2}) = D_1^W(V_{j=1}) \tag{33}$$

$*$ Find the normal points O_1 and P_2 on two neighboring axes 1 and 2 (figure 9):
With the given Z_1^W, Z_2^W, D_1^W, and D_2^W at configuration j=2 we obtain the following [26]: If $Z_1^W \times Z_2^W$ in figure 9 is not equal to zero, then

$$O_1^W = D_1^W + \frac{\left(D_2^W - D_1^W\right) \cdot \left(\left(Z_2^W \cdot Z_1^W\right) Z_2^W - Z_1^W\right)}{\left(Z_2^W \cdot Z_1^W\right)^2 - 1} Z_1^W \tag{34}$$

$$P_2^W = D_2^W + \frac{\left(D_2^W - D_1^W\right) \cdot \left(Z_2^W - \left(Z_2^W \cdot Z_1^W\right) Z_1^W\right)}{\left(Z_2^W \cdot Z_1^W\right)^2 - 1} \tag{35}$$

otherwise

$$O_1^W = D_1^W \tag{36}$$

$$P_2^W = D_2^W - Z_2^W \cdot \left(D_2^W - D_1^W\right) Z_2^W \tag{37}$$

$*$ Find a_1,α_1
Link parameters a_1 and α_1 can be found by

$$a_1 = \left|P_2^W - O_1^W\right| \tag{38}$$

and

$$\alpha_1 = \cos^{-1}\left(Z_1^W \cdot Z_2^W\right) \tag{39}$$

* Describe H_1^0

From equation (9), $a_0 = \alpha_0 = 0$, and the commanded joint reading v_1 at configuration j=2, we obtain

$$H_1^0 = \begin{bmatrix} c\theta_1 & -s\theta_1 & 0 & 0 \\ s\theta_1 & c\theta_1 & 0 & 0 \\ 0 & 0 & 0 & d_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{40}$$

where

$$\begin{aligned} \theta_1 &= s_1 \cdot v_1 \\ d_1 &= 0 \qquad for \;\; a \;\; revolute \;\; joint \end{aligned} \tag{41}$$

and

$$\begin{aligned} \theta_1 &= 0 \\ d_1 &= s_1 \cdot v_1 \qquad for \;\; a \;\; prismatic \;\; joint \end{aligned} \tag{42}$$

* Describe H_1^W

Link frame 1 relative to the world frame at configuration j=2 can be described as follows:
(1) Let the joint axis pose Z_1^W obtained at configuration j=2 be the Z axis of link frame 1.
(2) Let the point O_1^W be the origin of link frame 1.
(3) Let X_1^W of link frame 1 be described as

$$X_1^W = \frac{P_2^W - O_1^W}{|P_2^W - O_1^W|} \tag{43}$$

(4) The Y_1^W axis of link frame 1 is then found by taking the cross product of Z_1^W and X_1^W.

$$Y_1^W = Z_1^W \times X_1^W \tag{44}$$

(5) Finally, link frame 1 can be expressed as

$$H_1^W = \begin{bmatrix} X_1^W & Y_1^W & Z_1^W & O_1^W \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{45}$$

* Find the base-to-world transform H_B^W (or H_0^W):

From equations (40) and (45), H_B^W is obtained as

$$H_B^W = H_0^W = H_1^W \left(H_1^0\right)^{-1} \tag{46}$$

Case 2: $2 \leq i < m$

* Transform joint axes i-1, i, and i+1 to the same starting configuration j=i+1

Omit this step if the configurations for joint axes i-1, i, and i+1 are the same. In fact. we can omit this step if the values of v_1, v_2, ... v_{i-1} at configurations j=i-1 and j=i are equal to the values of v_1, v_2, ... v_{i-1} at configuration j=i+1. Otherwise, we do the following: Link frame i relative to the world frame at configuration j=i+1 can be represented as

$$H_i^W(V) = \begin{bmatrix} X_i^W & Y_i^W & Z_i^W & O_i^W \\ 0 & 0 & 0 & 1 \end{bmatrix} = H_0^W H_1^0(v_1) \ldots H_i^{i-1}(v_i) \tag{47}$$

Note that the above commanded joint readings V=(v_1, v_2, ... v_m) are recorded at configuration j=i+1. By inserting equations (9), (10), and (11) into equation (47) and with the given H_0^W, (a_1, ... a_{i-1}), (α_1, ... α_{i-1}), (d_1^*, ... d_{i-1}^*), (θ_1^*, ... θ_{i-1}^*), (s_1, ... s_{i-1}), and (v_1, ... v_{i-1}), we may easily obtain :

$$\begin{Bmatrix} Z_i^W \\ 0 \end{Bmatrix} = H_0^W * H_1^0 * \cdots * H_{i-1}^{i-2} * \begin{Bmatrix} 0 \\ -s\alpha_{i-1} \\ c\alpha_{i-1} \\ 0 \end{Bmatrix} \tag{48}$$

Replacing i by i-1 in equation (47), we also obtain the following:

$$\begin{Bmatrix} X_{i-1}^W \\ 0 \end{Bmatrix} = H_0^W * H_1^0 * \cdots * H_{i-2}^{i-3} * \begin{Bmatrix} c\theta_{i-1} \\ s\theta_{i-1} c\alpha_{i-2} \\ s\theta_{i-1} s\alpha_{i-2} \\ 0 \end{Bmatrix} \tag{49}$$

$$\begin{Bmatrix} O_{i-1}^W \\ 1 \end{Bmatrix} = H_0^W * H_1^0 * \cdots * H_{i-2}^{i-3} * \begin{Bmatrix} a_{i-2} \\ -s\alpha_{i-2} d_{i-1} \\ c\alpha_{i-2} d_{i-1} \\ 1 \end{Bmatrix} \tag{50}$$

From figures 9 and 7, we can see that

$$P_i^W = O_{i-1}^W + a_{i-1} X_{i-1}^W \tag{51}$$

Then, we select one reference point D_i coincident with the point P_i as

$$D_i^W = P_i^W \tag{52}$$

Therefore, equations (48) and (52) can be used to describe joint axis i at configuration j=i+1 .

* Find the normal points O_i and P_{i+1} on two neighboring axes i and i+1
With the given Z_i^W, Z_{i+1}^W, D_i^W, and D_{i+1}^W at configuration j=i+1 we obtain the following [26]: If Z_i^W x Z_{i+1}^W in figure 9 is not equal to zero, then

$$O_i^W = D_i^W + \frac{\left(D_{i+1}^W - D_i^W\right) \cdot \left(\left(Z_{i+1}^W \cdot Z_i^W\right) Z_{i+1}^W - Z_i^W\right)}{\left(Z_{i+1}^W \cdot Z_i^W\right)^2 - 1} Z_i^W \tag{53}$$

$$P_{i+1}^W = D_{i+1}^W + \frac{\left(D_{i+1}^W - D_i^W\right) \cdot \left(Z_{i+1}^W - \left(Z_{i+1}^W \cdot Z_i^W\right) Z_i^W\right)}{\left(Z_{i+1}^W \cdot Z_i^W\right)^2 - 1} \tag{54}$$

otherwise

$$O_i^W = D_i^W \tag{55}$$

$$P_{i+1}^W = D_{i+1}^W - \left(Z_{i+1}^W \cdot \left(D_{i+1}^W - D_i^W\right)\right) Z_{i+1}^W \tag{56}$$

$*$ Find a_i,α_i:
The link parameters a_i and α_i can be found by

$$a_i = \left|P_{i+1}^W - O_i^W\right| \tag{57}$$

and

$$\alpha_i = \cos^{-1}\left(Z_i^W \cdot Z_{i+1}^W\right) \tag{58}$$

$*$ Describe H_i^W
Link frame i relative to the world frame at configuration j=i+1 can be described as follows:
(1) Let the joint axis pose Z_i^W obtained at configuration j=i+1 be the Z axis of link frame i.
(2) Let the point O_i^W be the origin of link frame i.
(3) Let X_i^W of link frame i be described as

$$X_i^W = \frac{P_{i+1}^W - O_i^W}{|P_{i+1}^W - O_i^W|} \tag{59}$$

(4) The Y_i^W axis of link frame i is then found by taking the cross product of Z_i^W and X_i^W.

$$Y_i^W = Z_i^W \times X_i^W \tag{60}$$

(5) Finally, link frame i can be expressed as

$$H_i^W = \begin{bmatrix} X_i^W & Y_i^W & Z_i^W & O_i^W \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{61}$$

$*$ Find d_i^*, θ_i^* :
The parameters d_i^* and θ_i^* can be calculated by the following equations

$$\begin{aligned} \theta_i^* &= \theta_i - s_i \cdot v_i \\ d_i^* &= d_i \qquad for \quad a \quad revolute \quad joint \end{aligned} \tag{62}$$

where s_i is calculated from equation (28) and

$$d_i = \left|O_i^W - P_i^W\right| \tag{63}$$

$$\theta_i = \cos^{-1}\left(X_{i-1}^W \cdot X_i^W\right) \tag{64}$$

or

$$\begin{aligned} \theta_i^* &= \theta_i \\ d_i^* &= d_i - s_i \cdot v_i \qquad for \quad a \quad prismatic \quad joint \end{aligned} \tag{65}$$

where s_i is calculated from equation (30) and d_i and θ_i from equations (63) and (64), respectively .

Step 3: Representation of the total link transform H_m^B:
All constant parameters have been found as described in steps 1 and 2. Hence, the total link transform from the base frame B (or 0) to frame m can be represented in the following manner:

$$H_m^B(V) = H_1^0 * H_2^1 * \cdots * H_m^{m-1} \tag{66}$$

This means that the total link transform is a function of the commanded joint readings V=(v_1, v_2, ... v_m) read from the joint encoders.

Step 4: Calculation of the camera-to-manipulator transform H_c^m:
Link frame m at v_m=0 can be described as follows:
(1) Let the joint axis pose Z_m^W obtained in Step 1 be the Z axis of link frame m.
(2) Let the point P_m^W obtained in step 2 be the origin O_m^W of link frame m.
(3) Let X_{m-1}^W be the X axis (X_m^W) of link frame m.
(4) The Y axis of link frame m is then found by taking the cross product of Z_m^W and X_m^W.

$$Y_m^W = Z_m^W \times X_m^W \tag{67}$$

(5) Finally, the link frame can be expressed as

$$H_m^W = \begin{bmatrix} X_m^W & Y_m^W & Z_m^W & O_m^W \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{68}$$

In practice, the camera can be mounted on any link of the manipulator depending on how many degrees of freedom it needs. Therefore, we only need to calibrate up to the link m where the camera is mounted. The transform H_c^m can be found as

$$H_c^m = (H_W^c * H_m^W)^{-1} \tag{69}$$

Note that H_W^c is estimated at configuration j=m. For higher accuracy, many calibration configurations may be taken and the average of camera-to-manipulator matrices can be used.

Procedure for solving H_c^B(V) and H_W^c(V)

When the manipulator moves to a new position, the camera pose will be changed. The kinematic camera pose H_c^B which is a function of the manipulator's control variables (the commanded joint readings V) can be derived as follows:

$$H_c^B(V) = H_m^B(V) * H_c^m \tag{70}$$

Once the calibration procedure described above has been done, the camera pose can be obtained by inserting the current V into the above equation . No repeated calibration is

needed.
The camera pose H_W^c can be found in the same way using the following equation:

$$H_W^c(V) = (H_B^W * H_m^B(V) * H_c^m)^{-1} \tag{71}$$

Tsai's camera calibration

In order to avoid large scale nonlinear optimization, Tsai uses the radial alignment constraint(RAC) to decouple the calibration parameters into two groups. Then, the two groups are solved using the following two-stage approach:

Stage 1: The linear equation corresponding to each calibration point i is

$$[y_{di}X_{Wi} \quad y_{di}Y_{Wi} \quad y_{di} \quad -x_{di}X_{Wi} \quad -x_{di}Y_{Wi}] \cdot L \quad = \quad x_{di} \tag{72}$$

where L is a column vector containing the five unknowns $T_y^{-1}r_1$, $T_y^{-1}r_2$, $T_y^{-1}r_4$, $T_y^{-1}r_5$, and $T_y^{-1}T_x$. It takes a trick to solve R, T_x and T_y from L.

Stage 2: The linear equation to be solved in stage 2 is

$$\begin{aligned} H_{xi}f + H_{xi}R_d^2fk_1 - x_{di}T_z &= x_{di}(X_{Wi}r_7 + Y_{Wi}r_8) \\ H_{yi}f + H_{yi}R_d^2fk_1 - y_{di}T_z &= y_{di}(X_{Wi}r_7 + Y_{Wi}r_8) \end{aligned} \tag{73}$$

where

$$\begin{aligned} H_{xi} &= X_{Wi}r_1 + Y_{Wi}r_2 + T_x \\ H_{yi} &= X_{Wi}r_4 + Y_{Wi}r_5 + T_y \\ R_d^2 &= x_{di}^2 + y_{di}^2 \\ x_{di} &= s_x^{-1}(x_{fi} - C_x) \\ y_{di} &= s_y^{-1}(y_{fi} - C_y) \end{aligned}$$

3 EXPERIMENTS AND VERIFICATIONS

The goal of this section is to verify the procedure of obtaining H_c^B and H_W^c in Section **2.3**. Since it is impossible to measure the true H_c^B, we cannot compare the estimated and the true H_c^B directly. Instead, two indirect methods are used for the verification of H_c^B. The first method compares the measured and the predicted moving distance (or angle) while the second method compares the measured and the predicted 2D image data. In this section, the experimental setup is described first, then the experimental procedure of estimating H_c^B as a function of the joint readings V is elaborated. Following that, the two indirect methods are introduced.

The experimental setup consisted of a SONY XC-57 camera fastened to the last platform of a positioning table with two DOFs, one horizontal translation and one rotation about a vertical axis (see figure 10a). The focal length of the camera was 16mm. The calibration targets (see figure 10b) were an array of disks drawn on paper.

The experimental procedure for computing H_c^B and H_W^c is as follows:
For calibration of joint axis 1:
Step 1: Move the manipulator to the starting position for calibration. (A1 in figure 10c represents the camera center at the starting position and the arrow represents the camera optical axis.) Record the commanded joint readings V and estimate the camera pose H_W^c from the extrinsic camera calibration. Store the set of data (H_W^c, V). Note that the world coordinate system is defined by the calibration plate.
Step 2: Move that particular joint to the next position. (A2 in figure 10c is the camera center at the new position) After the movement, record the current commanded joint readings V and perform the extrinsic camera calibration. Store the new set of data (H_W^c, V).
Repeat steps 1 and 2 for joint axis 2. P1 and P2 in figure 10d represent the camera centers at the starting and new positions for joint axis 2. Then, we follow the procedures described in section 2.3 to obtain $H_m^B(\mathrm{V})$, $H_c^B(\mathrm{V})$, and $H_W^c(\mathrm{V})$. As indicated in equations (70) and (71), the kinematic camera pose H_c^B (or H_W^c) can be expressed in terms of V. Therefore with the given commanded joint readings V, the kinematic camera pose can be predicted.

Two indirect methods for the verification of H_c^B and H_W^c as follows:

Method 1:
Since the translation distance (or rotation angle) of a camera can be calculated from the estimated H_c^B, we can compare the estimated and the measured distance or angle to indirectly verify the estimated H_c^B. In the experiment, we move the camera to 11 positions along the translation axis and another 11 positions about the rotation axis . The commanded joint readings V at each position are recorded. Then, we substitute the readings V into the H_c^B function (found in the experimental procedure for H_c^B and H_W^c) to estimate H_c^B and then calculate the moving distances or angles . The true moving distances or angles for the positioning table were measured with high accuracy using encoders in the positioning table as a comparision. The result is listed in table 1 for translation and table 2 for rotation. The first row of tables 1 and 2 are the measured translation distance and rotation angle. The difference between the predicted and the measured values are listed in the second row. The third row is the percent error. Since all percent errors in tables 1 and 2 are less than 0.0081% and 0.373%, respectively, the proposed procedure for predicting H_c^B is quite accurate. The results in tables 1 and 2 also showed that the percentage error is decreased as the distance or rotation angle increases.

Method 2:
There are various ways using the function of H_c^B or H_W^c. One way is to predict the 2D image coordinates corresponding to 3D object points. The following experiment uses this

procedure to indirectly verify the function H_W^c. In the experiment, we move the camera to five positions (B1 through B5 in figure 10c) along axis 1 and five positions (Q1, Q2 in figure 10d and S1, S2, S3 in figure 10e) about axis 2. All the joint readings are recorded. For each position, substitute the joint readings into the function H_W^c. The 2D image coordinates can be predicted from the given intrinsic camera parameters and the known locations of the 3D target points. Since the 2D image data corresponding to the calibration target points can be directly measured at each camera position, the measured 2D image data are compared to the predicted data. The results for translation and rotation are listed in tables 3 and 4 respectively. Since 42 target points were used, the average errors listed in tables 3 and 4 are calculated from :

$$AverageError = \frac{\sum_{i=1}^{42}(meas.\ image\ coord.\ -\ predicted\ image\ coord.)}{42} \tag{74}$$

The result shows that average errors for translation range from 0.66 to 0.388 pixels and for rotation from 0.523 to 0.406 pixels. Note that the inaccuracy due to the prediction of H_W^c is contributing only part of the errors shown in tables 3 and 4. (The inaccuracies in the intrinsic camera parameters, the coordinates of the target points, the measured image coordinates are contributing errors to tables 3 and 4 as well.) The actual errors due to the prediction of H_W^c should be smaller than the errors listed in tables 3 and 4.

4 CONCLUSION

In this paper, we have presented a new technique for the unified calibration of an active camera system. In our approach, a Denavit- Hartenburg kinematic model, augmented with the capability of correcting errors caused by joint variable values used from joint readings, was established. Then, an analytical technique was developed to perform the manipulator calibration. Since the calibrated camera poses and link frames are related to the world frame, not only the camera-to-manipulator and base -to-world calibration is very straightforward, the procedure of solving

$$A\ H_c^m = H_c^m\ B$$

can also be eliminated. Two experiments were performed to verify the accuracy of the proposed technique. One was to examine the moving distances and angles, the other was to verify the estimated 2D image coordinates. Both experiments showed that the proposed technique is highly accurate.

The proposed method has the following advantages:
(1) Unified calibration - The calibration of the manipulator, camera-to-manipulator, camera, and base-to-world is integrated. The same experimental setup is used, and the intermediate results are shared.

(2) Simple computations - Only the form of one-equation-solving-one-parameter is used in the manipulator, camera-to-manipulator, and base-to-world calibration.
(3) Efficient - The camera-to-manipulator and base-to-world calibrations are the by-products of the manipulator calibration procedure.
(4) Accurate - For the manipulator, camera-to-manipulator, and base-to-world calibrations, the solution is in closed form and therefore accurate. No approximation is used.
(5) General use - A manipulator with any type of kinematic model can be used.
(6) No special equipment - Only a CCD camera and a calibration target are used.
(7) One-time calibration - Because the kinematic camera pose is expressed as a function of the joint readings, repeated calibration is not needed.

Acknowledgements

This work has been supported by the US Army Laboratory Command, Human Engineering Laboratory and Harry Diamond Laboratories. Special thanks go to Mr. Charles M. Shoemaker, formerly of the Human Engineering Laboratory and to Dr. Philip Emmerman of Harry Diamond Laboratories. The authors also appreciate the use of the source code for the camera calibration from Cheryl Sklair and William Hoff of Martin Marietta Space Systems.

References

[1] Abidi, M. A. and Chandra T. "Pose Estimation for Camera Calibration and Landmark Tracking", Proc. 1990 IEEE International Conference on Robotics and Automation.

[2] Aloimonos, J. and Badyopadhyay, A. "Active Vision", First Int'l Conf. on Computer Vision, IEEE, June 1987.

[3] Aloimonos, J. and Shulman, D., "Learning Shape Computations", Proc. Image Understanding Workshop, Feb. 1987.

[4] Aloimonos, J. "Purposive and Qualitative Active Vision", Proc. DARPA Image Understanding Workshop, 1990.

[5] Bajcsy, R. "Active Perception vs. Passive Perception", Proc. Workshop on Computer Vision, 1986.

[6] Bennett D. J., Hollerbach, J. M., and Geiger, D. "Autonomous Robot Calibration for Hand-eye Coordination" Robotics Research 5, 1990.

[7] Bottema, O. and Roth, B. "Theoretical kinematics" North-Holland publishing Co., 1979.

[8] Broderick, P. and Cipra, R. "A Method for Determining and Correcting Robot Position and Orientation Error Due to Manufacturing", Transaction of the ASME, J. of Mechanism, Transmissions, and Automation in Design, March 1988, Vol. 110.

[9] Chen H. H. "A Screw Motion Approach to Uniqueness Analysis of Head-Eye Geometry", Proc. CVPR, 1991.

[10] Chou, J. C. K. and Kamel, M. "Quaternions Approach to Solve the Kinematic Equation of Rotation,$A_a A_x = A_x A_b$, of a Sensor-mounted Robotic Manipulator", Proc. IEEE Int'l Conf. Robotics and Automation, 1988.

[11] Craig, J. "Introduction to Robotics: Mechanics and Control", Addison-Wesley publishing Co., Reading, MA, 1986.

[12] Denavit, J. and Hartenburg, R. S. "A Kinematic Notation for Lower-Pair Mechanisms Based on Matrices," ASME Journal of Applied Mechanics, June 1955.

[13] Hollerbach, J. M. "A Survey of Kinematic Calibration", The Robotic Review, The MIT Press, 1989.

[14] Hung, Y., Yeh. P., and Harwood, D., "Passive Ranging to Know Planar Point Sets," Proc. 1985 IEEE Int'l Conf. on Robotics and Automation.

[15] Lenz, R. K. and Tsai, R. Y. "Calibrating a Cartesian Robot with Eye-on-hand Configuation Independent of Eye-to-Hand Relationship" IEEE Transactions on pattern Analysis and Machine Intelligence, Vol. 11, No. 9, sep. 1989.

[16] Puskorius, G. V. and Feldkamp, L. A., "Global Calibration of a Robot/vision System," Proc. 1987 IEEE Int'l Conf. on Robotics and Automation.

[17] Raviv, D. and Herman, M., "Towards an Understanding of Camera Fixation," Proc. 1990 IEEE Int'l Conf. on Robotics and Automation.

[18] Shiu, Y. C. and Ahmad, S. "Calibration of Wrist-mounted Robotic Sensors by Solving Homogeneous Transform Equations of the Form AX=XB " IEEE Transactions on Robotics and Automation, Vol. 5, No. 1, Feb. 1989.

[19] Sklair, C., Gatrell, L., Hoff, W., and Magee, M., "Optical Target Location Using Machine Vision in Space Robotics Tasks," Proc. SPIE Symposium on Advances in Intelligent Systems Nov., 1990, SPIE Vol. 1387.

[20] Tsai, R. Y. "A Versative Camera Calibration Technique for High-accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV cameras and lenses " IEEE J. of Robotics and Automation, Vol. RA-3, No. 4, Aug. 1987.

[21] Tsai, R. Y. and Lenz, R. k. "Overview of a Unified Calibration Trio for Robot Eye, Eye-to-Hand, and Hand Calibration Using 3D Machine Vision" SPIE Vol. 1003 Sensor Fusion: Spatial Reasoning and Scene Interpretation, 1988.

[22] Tsai, R. Y. "Synopsis of Recent Progress on Camera Calibration for 3D Machine Vision" The Robotic Review, The MIT Press, 1989.

[23] Tsai, R. Y. and Lenz, R. K. "A New Technique for Fully Autonomous and Efficient 3D Robotics Hand/Eye Calibration" IEEE Transactions on Robotics and Automation, Vol. 5, No. 3, June 1989.

[24] Vernon, D. and Tistarelli, M. "Using camera motion to estimate range for robotic parts manipulation" IEEE Transactions on Robotics and Automation, Vol. 6, No. 5, Oct. 1990.

[25] Zheng, J. Y., Chen, Q., and Tsuji, S. "Active Camera Guided Manipulation," Proc. 1991 IEEE Int'l Conf. on Robotics and Automation.

[26] Zhuang, H., Roth, Z., and Hamano, F. "A Complete and parametrically Continuous Kinematic Model for Robot Manipulator" Proc. 1990 IEEE Int'l Conf. on Robotics and Automation.

[27] Ziegert, J. and Datseris, P. "Robot Calibration Using Local Pose Measurements" Int'l J. of Robotics and Automation, Vol. 5, No. 2, 1990.

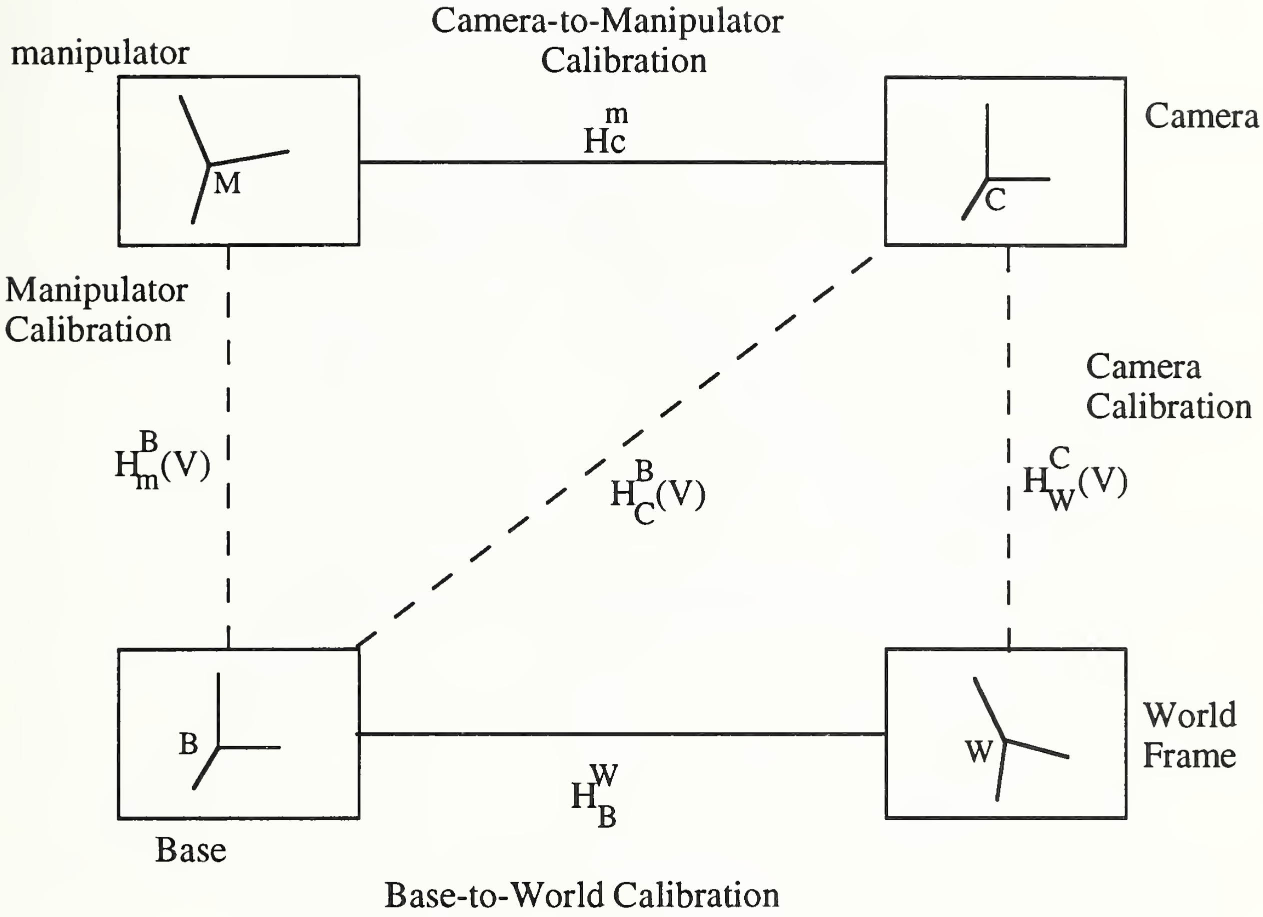

Camera-to-Manipulator Calibration
manipulator
M
H_c^m
Camera
C
Manipulator Calibration
Camera Calibration
$H_m^B(V)$
$H_C^B(V)$
$H_W^C(V)$
B
W
World Frame
H_B^W
Base
Base-to-World Calibration

Fig. 1 Unified Calibration

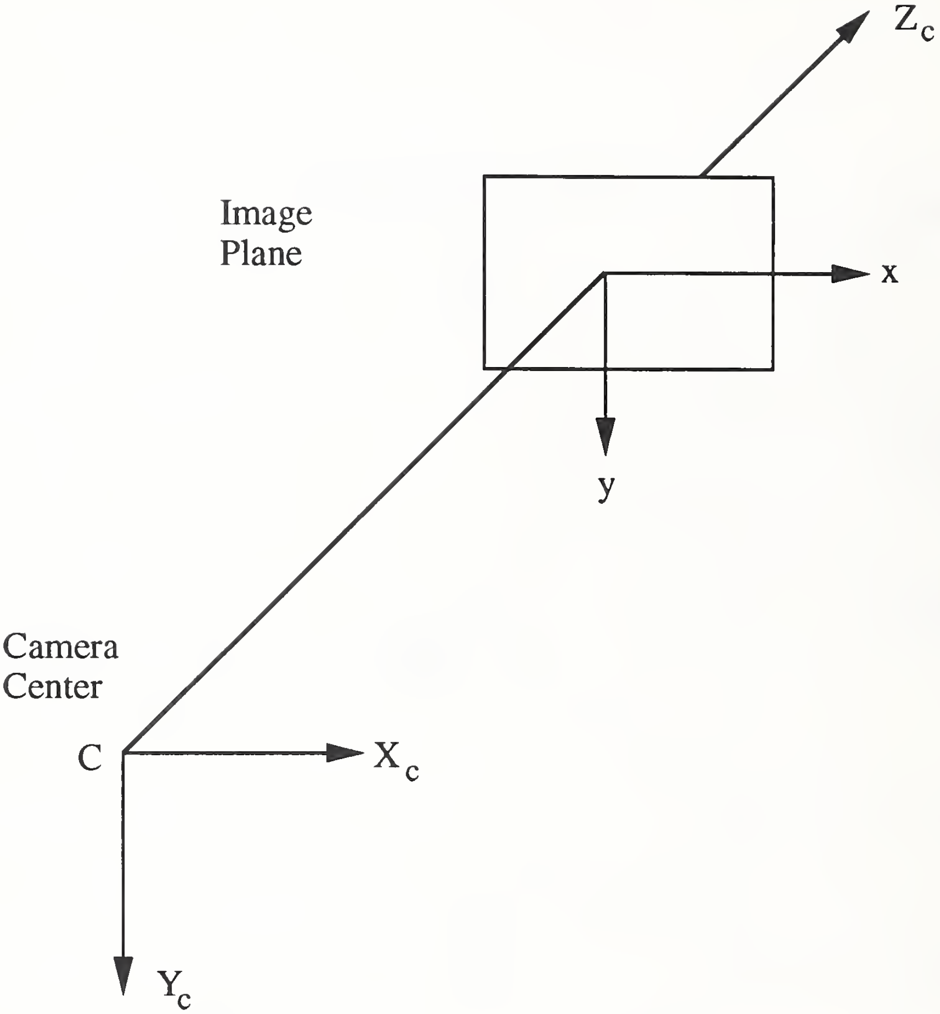

Z_c
Image
Plane
x
y
Camera
Center
C
X_c
Y_c

Fig. 2 Camera Coordinate Frame

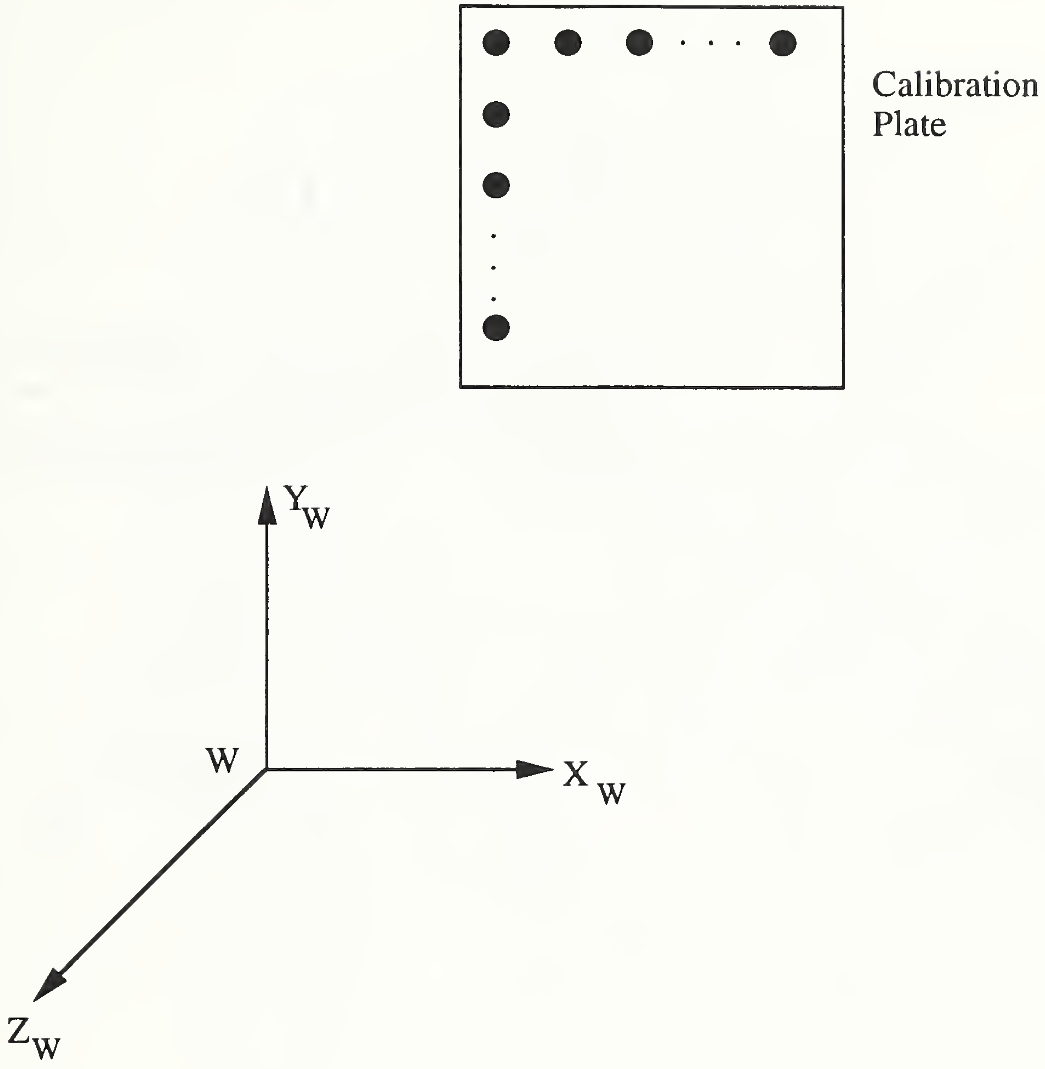

Calibration Plate
Y_W
W
X_W
Z_W

Fig. 3 World Coordinate Frame

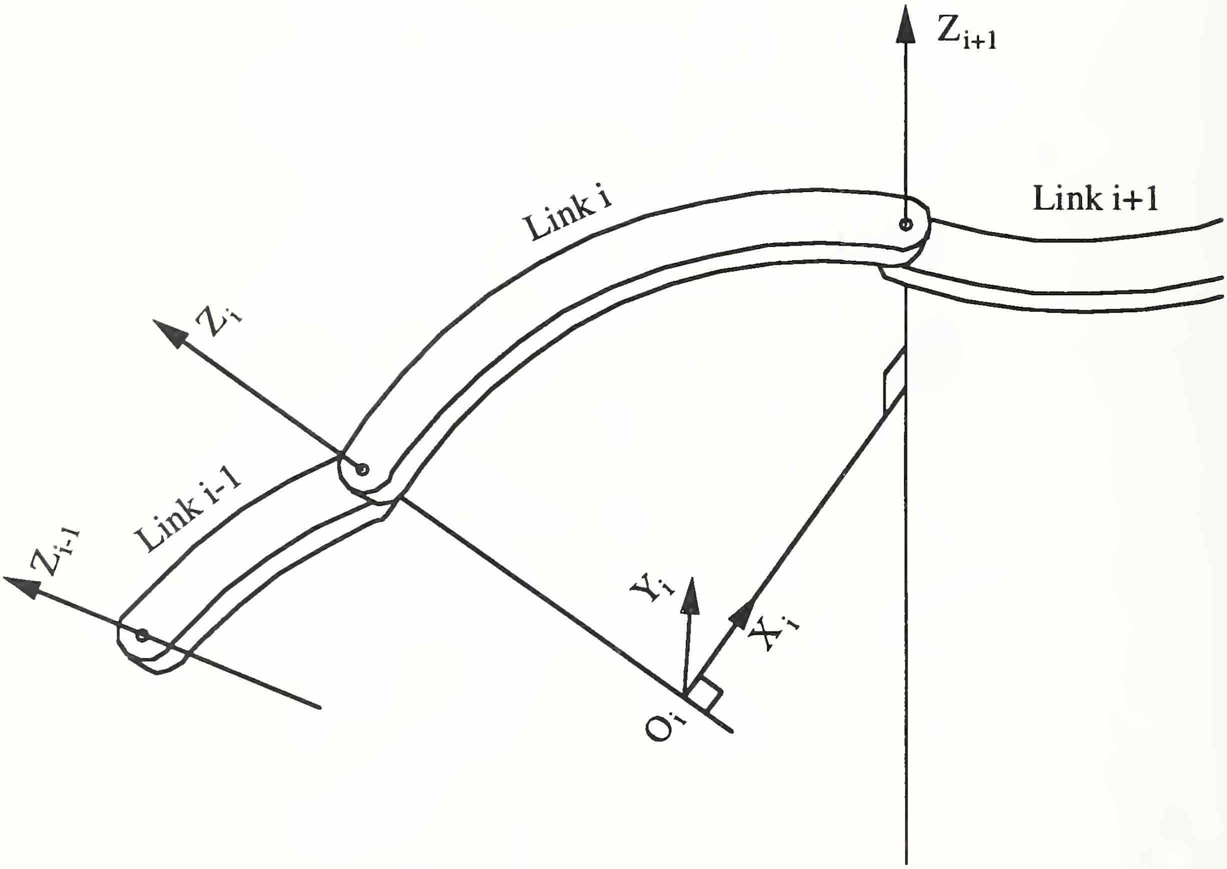

Z_{i+1}
Link i+1
Link i
Z_i
Link i-1
Z_{i-1}
Y_i
X_i
O_i

Fig. 4 Link Coordinate Frame

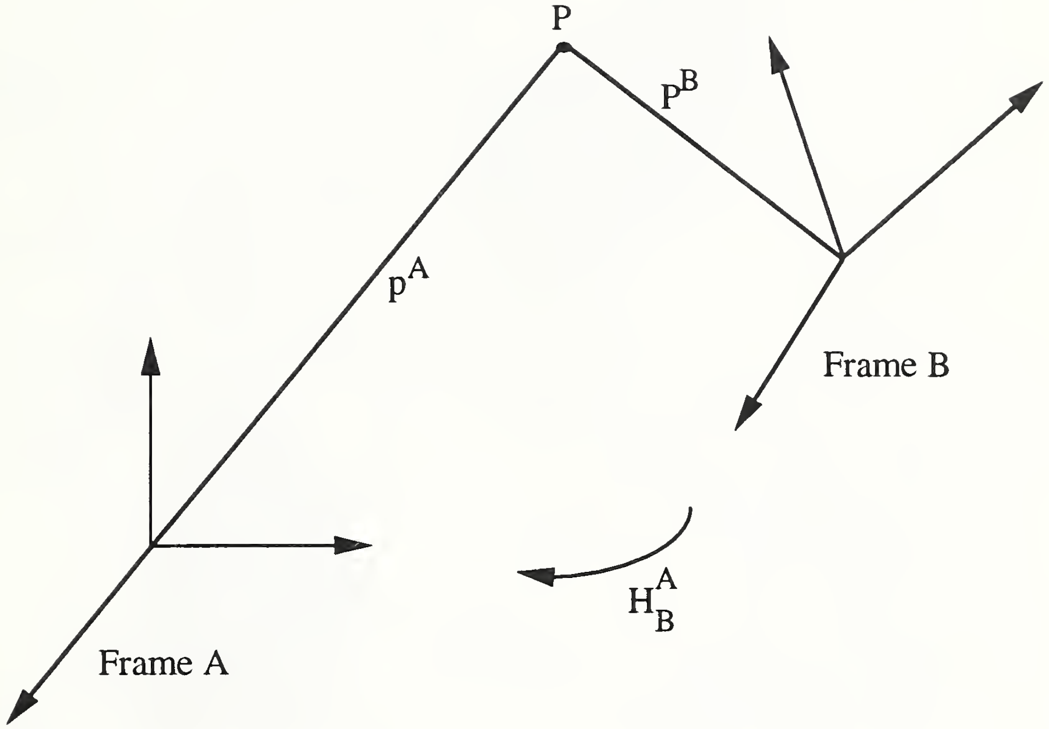

P
p^B
p^A
Frame B
Frame A
H_B^A

Fig. 5 Transform Mapping

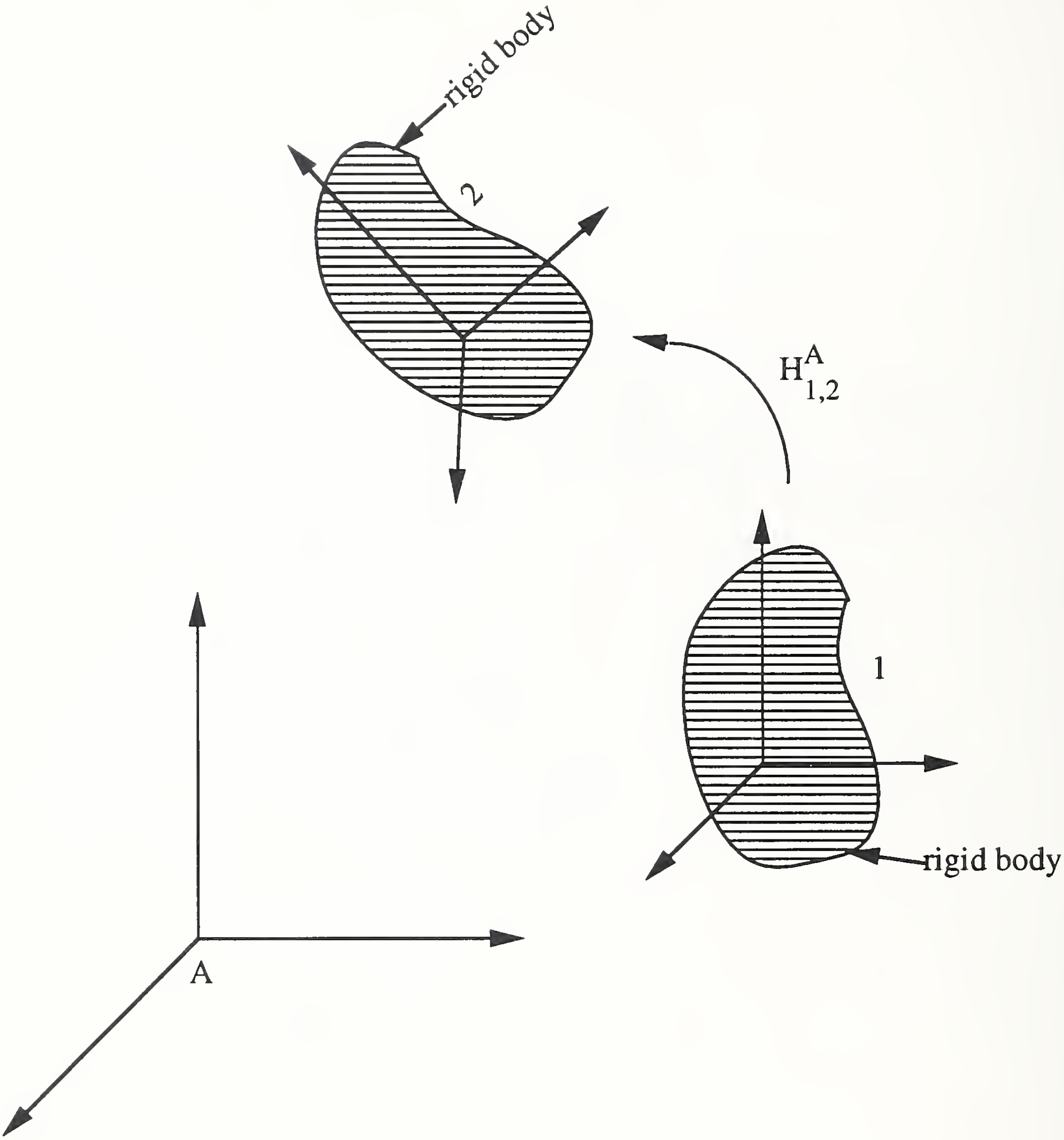

rigid body
2
$H^{A}_{1,2}$
1
rigid body
A

Fig. 6 Transform Operator

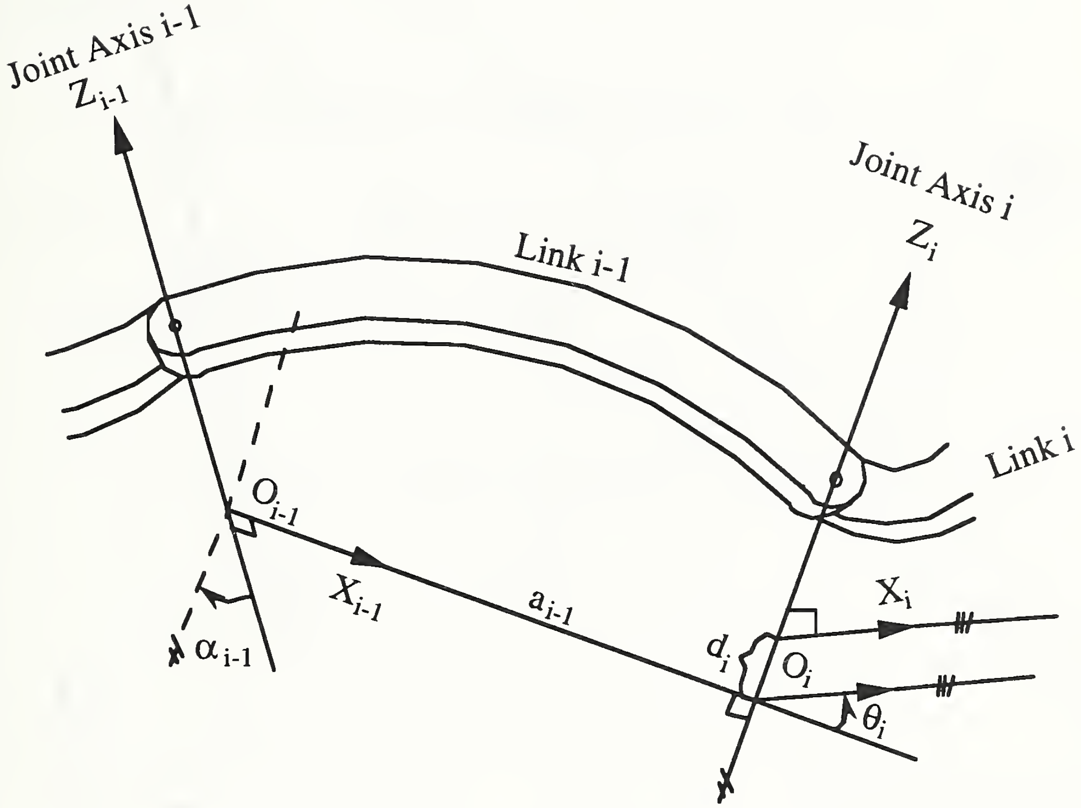
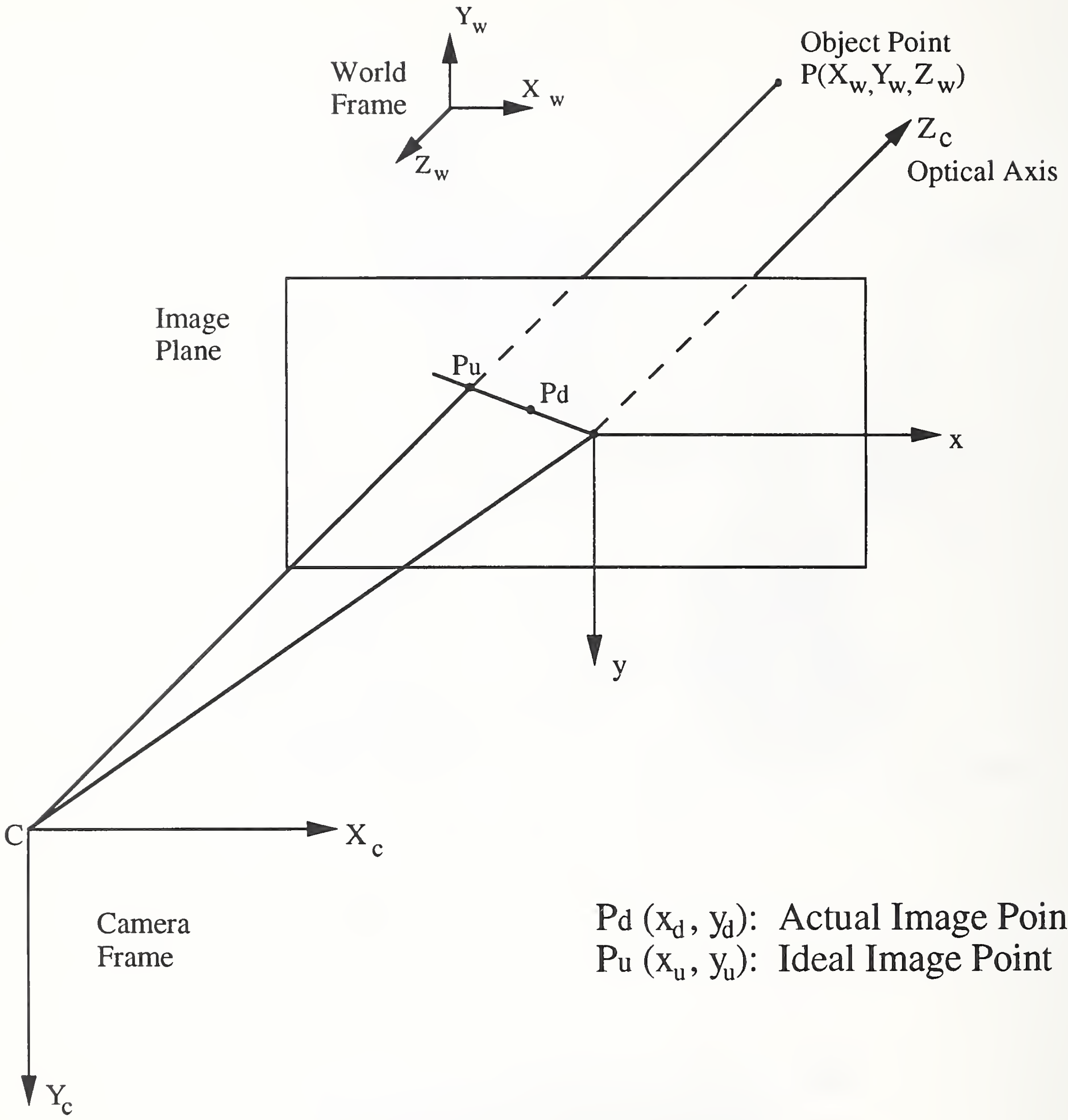

Joint Axis i-1
Z_{i-1}
Link i-1
Joint Axis i
Z_i
Link i
O_{i-1}
X_{i-1}
a_{i-1}
α_{i-1}
X_i
d_i
O_i
θ_i

Fig. 7 Link Coordinate Frames and Parameters of a Manipulator


Y_w
World Frame
X_w
Z_w
Object Point
P(X_w, Y_w, Z_w)
Z_c
Optical Axis
Image Plane
Pu
Pd
x
y
C
X_c
Camera Frame
Y_c

Pd (x_d, y_d): Actual Image Point
Pu (x_u, y_u): Ideal Image Point

Fig. 8 Camera Model

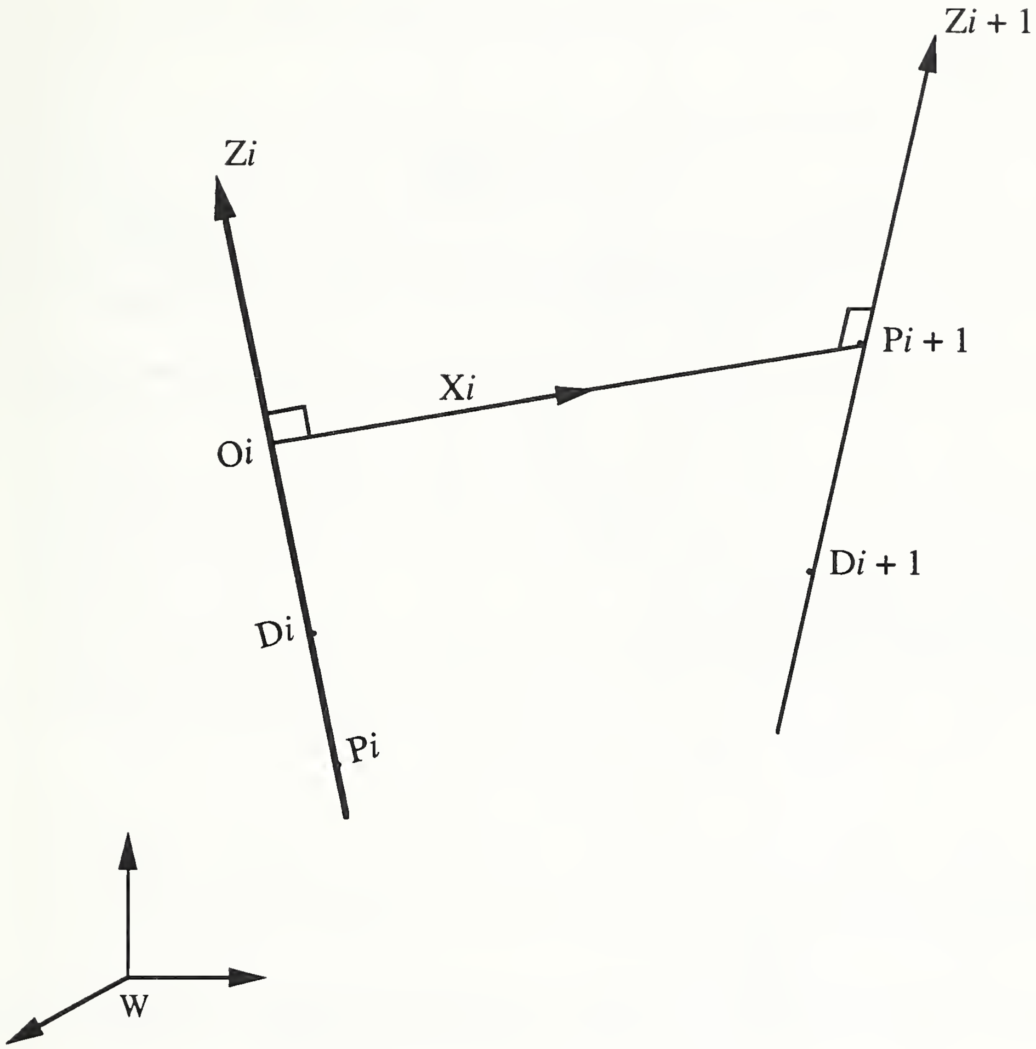

Zi + 1
Zi
Pi + 1
Xi
Oi
Di + 1
Di
Pi
W

Fig. 9 The Intersection Line Perpendicular to the Two Joint Axes

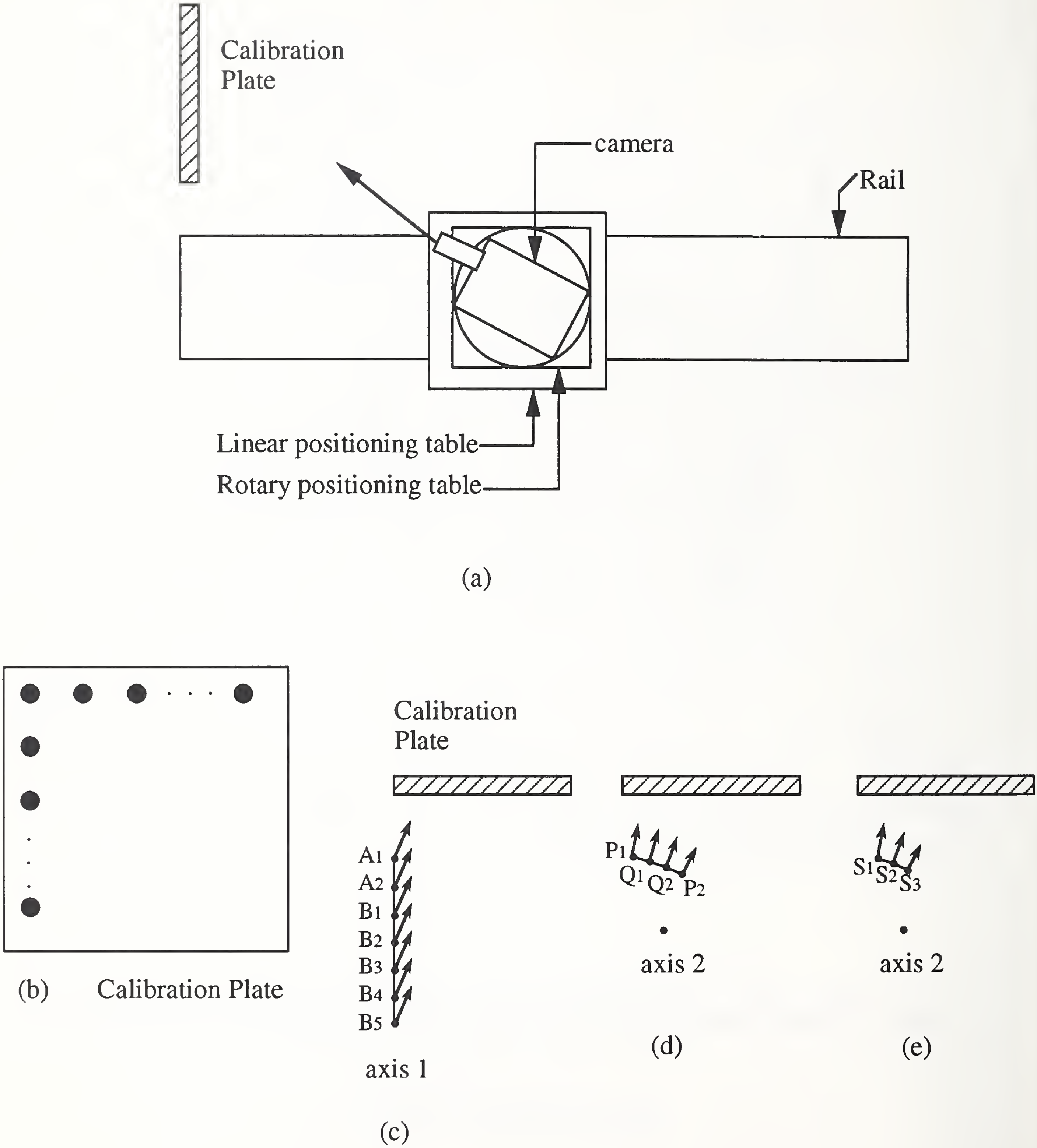

Calibration Plate
camera
Rail
Linear positioning table
Rotary positioning table
(a)
(b) Calibration Plate
Calibration Plate
A1
A2
B1
B2
B3
B4
B5
axis 1
(c)
P1
Q1
Q2
P2
axis 2
(d)
S1
S2
S3
axis 2
(e)

Fig. 10 Experimental Setup

Translation (mm)	100	200	300	400	500	600	700	800	900	1000
error (mm)	0.0081	0.013	0.018	0.023	0.028	0.033	0.039	0.044	0.049	0.054
error (%)	0.0081	0.0066	0.0061	0.0059	0.0057	0.0056	0.0055	0.0055	0.0054	0.0054

Table 1: Results of experiment in method 1 (translation)

Rotation Angle (rad)	0.052	0.104	0.157	0.209	0.262	0.349	0.523	0.698	0.872	1.047
error (10^{-3} rad)	0.20	0.31	0.44	0.58	0.73	0.97	1.5	1.9	2.4	2.9
error (%)	0.373	0.294	0.282	0.278	0.277	0.277	0.277	0.277	0.278	0.278

Table 2: Results of experiment in method 1 (rotation)

Camera Position	B1	B2	B3	B4	B5
Average Error (pixel)	0.660	0.614	0.523	0.433	0.388

Table 3: Results of experiment in method 2 (translation)

Camera Position	Q1	Q2	S1	S2	S3
Average Error (pixel)	0.513	0.433	0.523	0.466	0.406

Table 4: Results of experiment in method 2 (rotation)

NIST-114A
(REV. 3-90)

U.S. DEPARTMENT OF COMMERCE
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY

BIBLIOGRAPHIC DATA SHEET

1. PUBLICATION OR REPORT NUMBER
NISTIR 4715

2. PERFORMING ORGANIZATION REPORT NUMBER

3. PUBLICATION DATE
NOVEMBER 1991

4. TITLE AND SUBTITLE

Kinematic Calibration of an Active Camera System

5. AUTHOR(S)

Gin-Shu Young, Tsai-Hong Hong, Martin Herman, and Jackson C.S. Yang

6. PERFORMING ORGANIZATION (IF JOINT OR OTHER THAN NIST, SEE INSTRUCTIONS)
U.S. DEPARTMENT OF COMMERCE
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
GAITHERSBURG, MD 20899

7. CONTRACT/GRANT NUMBER

8. TYPE OF REPORT AND PERIOD COVERED

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (STREET, CITY, STATE, ZIP)

10. SUPPLEMENTARY NOTES

11. ABSTRACT (A 200-WORD OR LESS FACTUAL SUMMARY OF MOST SIGNIFICANT INFORMATION. IF DOCUMENT INCLUDES A SIGNIFICANT BIBLIOGRAPHY OR LITERATURE SURVEY, MENTION IT HERE.)

Perceptual activity for exploration, probing and searching is very important in computer vision. For the purpose of intelligently controlling the sensor's motion and parameters for different sensing strategies and designated tasks in perceptual activity , an active camera system is often used. An active camera system is also usually used for vision-based guidance. In order to position the camera system accurately and to obtain the relation between the camera and the manipulator, the active camera system must be calibrated.

In this paper, we introduce a modified Denavit-Hartenburg kinematic model and develop a new technique to calibrate an active camera system. The manipulator, camera-to-manipulator, camera, and base-to-world calibrations are all included. The method employs four ideas : (1)The camera poses, joint poses, and link frames calibrated are all related to the world frame, therefore the camera-to-manipulator and base-to-world calibration is very straightforward; (2)The joint poses are calibrated separately; (3)The manipulator motion is obtained from the camera poses; (4)Once the joint poses are obtained, the link frames can be defined for any kinematic model. In consequence, the whole procedure is simple, flexible, accurate, and efficient. Two experiments are performed to verify the accuracy of the new technique.

12. KEY WORDS (6 TO 12 ENTRIES; ALPHABETICAL ORDER; CAPITALIZE ONLY PROPER NAMES; AND SEPARATE KEY WORDS BY SEMICOLONS)
active camera calibration; active vision; camera-to-manipulator calibration; manipulator calibration; static camera calibration; visual motion

13. AVAILABILITY
☒ UNLIMITED
☐ FOR OFFICIAL DISTRIBUTION. DO NOT RELEASE TO NATIONAL TECHNICAL INFORMATION SERVICE (NTIS).
☐ ORDER FROM SUPERINTENDENT OF DOCUMENTS, U.S. GOVERNMENT PRINTING OFFICE, WASHINGTON, DC 20402.
☒ ORDER FROM NATIONAL TECHNICAL INFORMATION SERVICE (NTIS), SPRINGFIELD, VA 22161.

14. NUMBER OF PRINTED PAGES
36

15. PRICE
A03

ELECTRONIC FORM